Three- and More-way Tables

		Page
Introduction		14-1
14A.	A psychological experiment: first step	14-3
14B.	Single-factor elementary analyses	14-6
14C.	Moving onwards	14-18
14D.	The ell and the block	14-24
14E.	Choice of expression	14-29
14F.	Inconsistent factors	14-47
14G.	Sweeping out with elementary analyses	14-56
14H.	How far have we come?	14-69

*****

14K.	One way to add resistance	14-70

Location of tables and plots for 14

Number	Page	Number	Page
1	14-4	26	14-45
2	14-7	27	14-46
3	14-8	28	14-49
4	14-9	29	14-51
5	14-11	30	14-52
6	14-12	31	14-54
7	14-13	32	14-55
8	14-15	33	14-59
9	14-16	34	14-60
10	14-17	35	14-61
11	14-20	36	14-62
12	14-22	37	14-63
13	14-23	38	14-64
14	14-25	39	14-65
15	14-27	40	14-66
16	14-28	41	14-67
17	14-31	42	14-68
18	14-33	43	14-73
19	14-34	44	14-74
20	14-35	45	14-75
21	14-37	46	14-77
22	14-38		
23	14-40		
24	14-41		
25	14-44		

14

Three- and More-way Tables

Tables can easily involve three or more factors, setting out responses for all combinations of certain versions of each factor:

- temperatures can involve a month like May, a year like 1959, and a place like Washington.
- voting behavior can involve any combination of one of several counties, one of several elections, and one of several offices.
- human weights can involve any combination of one of several ages, one of two sexes, and one of several diets.
- recipes almost always involve several independently variable ingredients; one cake-mix experiment involved ten such factors.
- most industrial processes involve at least three, usually many more, choices that can be made separately.

With what we already understand about two-way tables, we can go much further in dealing with three- and more-way tables than we might expect following the familiar process of breaking each value up into a sum of terms.

We break up one-way tables -- batches -- into the sum of two terms: typical and residuals.

We break up two-way tables into the sum of two times two -- four -- terms: common, row adjustments, column adjustments, residuals.

We are now going to learn about breaking up three-way tables into the sum of two times two times two -- eight -- terms: seven terms of description, one of residuals.

Once we have done this, we can see our way to breaking up an n-way table into 2^n terms: 2^n-1 terms of description, one of residuals.

We will also learn how to stop at various intermediate stages.

The possibility of doing all this, and the general character of what is done, does not depend on how we choose to fit. We can do it using means -- or midmean differences -- or streamlined procedures. Resistance is still important; indeed, cleanliness of analysis can only become even more important as data becomes organized in more detail and in more ways. Relative ease of analysis is also steadily more important as the data becomes more highly organized. Finding a method of bearable computational difficulty that has most of the advantages of a full use of midmean differences is likely to be of extreme practical importance.

A good method of tearing the data apart is of little use to us, however, if we do not understand the terms into which it has been torn, or the way in which they fit together. Our first effort has to be understanding. Gaining this for three- and more-way tables is difficult enough to demand keeping everything as simple as possible.

Analysis by means has great advantages of simplicity: Doing things in different orders gives identically the same results -- except for the effects of keeping only a limited

number of decimal places. It is always easy to check that the mean has been taken out exactly -- the sum of what remains ought to be zero. Even typical squares add up -- exactly. Accordingly, even though the results thus found are not as resistant as they might be, the body of this chapter will be confined to analysis by means.

14A. A psychological experiment: first step

In 1944, Johnson and Tsao published an account of an experiment for measuring the sensitivity of individuals to changes in pull on a ring held by one finger. Measurements were made with different initial steady pulls (here called weight) and different rates of increase of the pull once increase was begun (here called rate). The persons tested included four who were blind and four who could see. The measurements recorded were the increase in pull before the subjects reported feeling a change.

A small part of their data provides an example of convenient size, one that can show us how two-way methods can extend to three-way data. Table **1** gives the sums, over two experimental sessions, called "dates" by Johnson and Tsao, of the mean of five responses on each date. (The values given have been rounded to whole numbers for simplicity; they are expressed in grams of weight.) Table 1 involves:

- three persons, all blind, called here **K**, **L** and **M** (IB(2), IIB(1), and IIB(2) in Johnson's notation).
- three initial pulls (weights), called here **1**, **4**, and **7** (as in Johnson 1949) and amounting to 100, 250 and 400 grams.

table 1 of chapter 14

The 3x3x4 = 36 Basic Observations and 3x4 = 12 Means

(Johnson and Tsao example)

Rate & Subject	Initial Pull (Weight) 1	4	7	(Total)	Mean
aK	39	31	26	(96)	32.0
aL	16	12	12	(40)	13.3
aM	18	13	14	(45)	15.0
bK	85	55	56	(196)	65.3
bL	32	22	20	(74)	24.7
bM	31	26	26	(83)	27.7
cK	101	84	70	(255)	85.0
cL	43	33	30	(106)	35.3
cM	42	38	40	(120)	40.0
dK	151	124	98	(373)	124.3
dL	63	51	37	(151)	50.3
dM	58	55	46	(159)	53.0

Source: Palmer O. Johnson 1949. Statistical Methods in Research, Table 87 on page 290.

- four rates of increase, called here **a**, **b**, **c**, and **d** (as in Johnson 1949), amounting to 100, 200, 300 and 400 grams per minute.

This table, as is almost necessary for a table put on a page, is arranged in rows and columns. The columns correspond to one variable (weight), while the rows correspond to two variables (subject and rate). In any situation involving all combinations of certain versions of three or more variables we can think of the numbers laid out similarly, with certain variables combined into columns and others combined into rows. In Johnson's full table of 448 numbers (1949, p. 290) for example:

- sex, sight, date and individual are combined into columns, of which there are $2x2x2x2 = 16$.

- weight and rate are combined into rows, of which there are $7x4 = 28$.

Arrangements of this kind are the only natural way to put the numbers down. They also provide, as we shall see, a convenient way to begin their analysis.

Where we have 3,4,...,10,... factors, the problem of tagging them can be somewhat troublesome. Surely "row" and "column" are no longer useful. Two choices are natural:

- when we are analyzing "for real", tags linked to the way we think of the factor -- such as "rate" for rate, "init" for initial pull, "subj" for person who is the subject -- are likely to be best.

- when, as in the next few sections, we are analyzing to learn about analysis, it often helps to link the tag for a factor to the tags for its versions -- as when **KLM** is used to tag the factor whose versions are **K**, **L**, or **M**.

We shall follow the latter course for the present.

Returning to table 1, where row totals and row means are given, we can subtract the row means, forming the first residuals shown in table **2** (with column totals and column means). Removing column means takes us to the second residuals of table **3** and completes the elementary analysis of table 1, regarded as a two-way table.

If we had really had only a two-way table, table 3 would have left us pausing to reflect as to how we might go somewhat further. As a part of such a pause, we would perhaps have looked at a stem-and-leaf display for the residuals, like that of table **4**. The pattern of this display, with roughly 1/4 of all values separated and high and closely another 1/4 separated and low, certainly furnishes ground for belief that further serious analysis remains to be done. (We will come back to these residuals in section 16E.)

14B. Single-factor elementary analyses

If we are to understand how to analyze three-way tables, it can help to make up some and then see what happens as we try analyses. As a simple beginning, then, let us set up a table -- with the same size, shape, and labels as the Johnson and Tsao example -- each cell of which can be written as a sum of four terms:

- for every cell a common 1000.
- for every cell a term depending on whether it is **a**, **b**, **c**, or **d** -- namely -40, -10, 10, or 40, respectively.
- for every cell a term depending on whether it is **K**, **L**, or **M** -- namely, 300, -200, or -100, respectively.
- for every cell a term depending on whether it is **1**, **4**, or **7** -- 9, -2, or -7, respectively.

table 2 of chapter 14

First Residuals Based on table 1

Rate & Subject	Initial Pull (Weight)		
	1	4	7
aK	7.0	-1.0	-6.0
aL	2.7	-1.3	-1.3
aM	3.0	-2.0	-1.0
bK	19.7	-10.3	-9.3
bL	7.3	-2.7	-4.7
bM	3.3	-1.7	-1.7
cK	16.0	-1.0	-15.0
cL	7.7	-2.3	-5.3
cM	2.0	-2.0	0.0
dK	26.7	-0.3	-26.3
dL	12.7	0.7	-13.3
dM	5.0	2.0	-7.0
(Total)	(113.1)	(-21.9)	(-90.9)
Mean	9.4	-1.8	-7.6

table 3 of chapter 14

Second Residuals Based on table 2
as Part of Elementary Analysis of table 1

Rate & Subject	Initial Pull (Weight) 1	4	7	
aK	-2.4	0.8	1.6	-15.2
aL	-6.7	0.5	6.3	-33.9
aM	-6.4	-0.2	6.6	-32.2
bK	10.3	-8.5	-1.7	18.1
bL	-2.1	-0.9	2.9	-22.5
bM	-6.1	0.1	5.9	-19.5
cK	6.6	0.8	-7.4	37.8
cL	-1.7	-0.5	2.3	-11.9
cM	-7.4	-0.2	7.6	-7.2
dK	17.3	1.5	-18.7	77.1
dL	3.3	2.5	-5.7	3.1
dM	-4.4	3.8	0.6	5.8
	9.4	-1.8	-7.6	47.2

table 4 of chapter 14

Stem-and-leaf Distribution for the 36 Second Residuals of table 3 (also hinges and rough straggler check)

2						
1	07				07	
9						
8						
7	6				6	
6	366				366	
5	9				9	
4						
3	38				38	
				H		
2	593		369			
1	56		66			
+0	85186		15688			
-0	2952		9522			
-1	77		77			
-2	14		41			
-3		H				
-4	4	4				
-5	7	7				
-6	147	741				
-7	44	44				
-8	4	4				
-9						
-1	8	8				
-2						

Hinges: +3.1
-3.4

Midspread: 6.5

Cuts: 9.6
-9.9

Notes: 10 is 10.3, 17 is 17.3, -18 is -18.7.

These are simple constituents simply assembled. A satisfactory analysis ought to bring out their essence, though not necessarily their exact detail.

Table **5** shows the result of putting these terms together. Table **6** shows the result of a two-way table elementary analysis carried to the point to which we have carried the Johnson data. The results are those which we might have expected, had we stopped to think: the common term and the **147** terms have been neatly separated into separate cells, while the **abcd** and **KLM** terms are still entangled with one another. Clearly a further step is required.

The rows of tables 5 and 6 -- like those of tables 1, 2, and 3 -- have a compound label. Any set of numbers tagged by these rows can appropriately be written as a two-way table. In particular, we can do this with the row adjustments -- 260,-240,-140,290, ... -- of the elementary analysis of table 6. Once we do this, as we do in table **7**, it is natural to go on to a further elementary analysis.

This latter elementary analysis neatly separates the two still-entangled terms, as we knew it would.

In panel **c**, we have shown the last elementary analysis in an unfamiliar form. Because we are using compound row labels, the data -- and each of the four terms into which it is analyzed -- is rather naturally written in a single column.

In this table, we have used a pattern of "**bold** on first appearance" to link labels and terms. Thus the

table **5** of chapter 14

A Synthetic Example of a 3x3x4 Table
(To be called example A)

a) THE TABLE

	(*)	**1**	**4**	**7**
(**)		(1009)	(998)	(993)
aK	(260)	1269	1258	1253
aL	(-240)	769	758	753
aM	(-140)	869	858	853
bK	(290)	1299	1288	1283
bL	(-210)	799	788	783
bM	(-110)	899	888	883
cK	(310)	1319	1308	1303
cL	(-190)	819	808	803
cM	(-90)	919	908	903
dK	(340)	1349	1338	1333
dL	(-160)	849	838	833
dM	(-60)	949	938	933

(*) = (sum of **abcd** and **KLM** terms).

(**) = (sum of **common** and **147** terms).

b) THE CONSTITUENT TERMS:

common: 1000

abcd: **a** = -40, **b** = -10, **c** = 10, **d** = 40.

KLM: **K** = 300, **L** = -200, **M** = -100.

147: **1** = 9, **4** = -2, **7** = -7.

c) EXAMPLES: 260 = -40+300, 1009 = 1000+9, 1269 = 260+1009.

table 6 of chapter 14

Two-way Elementary Analysis of Example A

	1	4	7	
aK	0	0	0	260
aL	0	0	0	-240
aM	0	0	0	-140
bK	0	0	0	290
bL	0	0	0	-210
bM	0	0	0	-110
cK	0	0	0	310
cL	0	0	0	-190
cM	0	0	0	-90
dK	0	0	0	340
dL	0	0	0	-160
dM	0	0	0	-60
	9	-2	-7	1000

table 7 of chapter 14

The Row Adjustments of table 6 Themselves Analyzed
(Example A)

a) ROW ADJUSTMENTS IN TWO-WAY FORM:

	a	b	c	d
K	260	290	310	340
L	-240	-210	-190	-160
M	-140	-110	-90	-60

b) ELEMENTARY ANALYSIS OF SAME:

	a	b	c	d	
K	0	0	0	0	300
L	0	0	0	0	-200
M	0	0	0	0	-100
	-40	-10	10	40	0

c) THE SAME ELEMENTARY ANALYSIS IN COMPOUND-ROW FORM

	The data	The elementary analysis			
		(xx)	(x)	(x)	(·)
aK	260	0	-40	300	0
aL	-240	0	-40	-200	
aM	-140	0	-40	-100	
bK	290	0	-10	300	
bL	-210	0	-10	-200	
bM	-110	0	-10	-100	
cK	310	0	10	300	
cL	-190	0	10	-200	
cM	-90	0	10	-100	
dK	340	0	40	300	
dL	-160	0	40	-200	
dM	-60	0	40	-100	

Note: (·) = common term
(x) = single factor term
(xx) = two factor term

appearance of **300, -200, -100** in the next to the last column matches the **K, L, M** -- given as a**K**, a**L**, a**M** -- in the stub. We shall continue to use this sort of "linkage-by-**bolding**" , as well as the notation "(·)" for common "(x)" for single-factor terms, "(xx)" for two-factor terms, and so on. The use of a vertical bar to replace a long segment of repetitions will also be standard.

If we join together both the styles and the analyses of tables 6 and 7, we obtain the result of table **8**. Here we have an analysis of example A into five terms:

- residual.
- common.
- three adjustments, one for each separate factor.

This is not yet the eight-term break we promised, but it is a long step toward it.

* * *

Having got this far with example A, it is time to bring the Johnson and Tsao example up to date. Table **9** shows the elementary analysis of the (compound) row adjustments. Table **10** shows the result of combining the analyses of tables 3 and 9. In combining these results, there are two separately obtained terms that have to be combined in order to form the "residuals" , these are:

- the residuals of table 3.
- the residuals of table 9, now best thought of as the last column of panel 3**c**.

Neither of these terms are taken care of by any of the single-factor adjustments set out in table 10. Accordingly

table 8 of chapter 14

One-factor Three-way Elementary Analysis of table 5
(Example A)

	1	4	7	(x)	(x)	(ø)
aK	0	0	0	-40	300	1000
aL	0	0	0	-40	-200	
aM	0	0	0	-40	-100	
bK	0	0	0	-10	300	
bL	0	0	0	-10	-200	
bM	0	0	0	-10	-100	
cK	0	0	0	10	300	
cL	0	0	0	10	-200	
cM	0	0	0	10	-100	
dK	0	0	0	40	300	
dL	0	0	0	40	-200	
dM	0	0	0	40	-100	
(x)	9	-2	-7			

table 9 of chapter 14

Elementary Analysis of the (Compound) Row Adjustments
found in table 3
(Johnson and Tsao example)

a) ROW ADJUSTMENTS IN TWO-WAY FORM:

	a	**b**	**c**	**d**
K	-15.2	18.1	37.8	77.1
L	-33.9	-22.5	-11.9	3.1
M	-32.2	-19.5	-7.2	5.8

b) ELEMENTARY ANALYSIS FOR SAME:

	a	**b**	**c**	**d**	
K	-17.6	-3.4	2.1	18.9	29.5
L	9.5	1.8	-1.8	-9.3	-16.3
M	8.2	1.8	-0.1	-9.6	-13.3
	-27.1	-8.0	6.2	28.7	0

c) SAME INPUT AND ELEMENTARY ANALYSIS USING COMPOUND ROWS

	(Input)	(xx)	(x)	(x)	(·)
aK	-15.2	-17.6	**-27.1**	**29.5**	0.0
a**L**	-33.9	9.5	-27.1	**-16.3**	
a**M**	-32.2	8.2	-27.1	**-13.3**	
bK	18.1	-3.4	**-8.0**	29.5	
bL	-22.5	1.8	-8.0	-16.3	
bM	-19.5	1.8	-8.0	-13.3	
cK	37.8	2.1	**6.2**	29.5	
cL	-11.9	-1.8	6.2	-16.3	
cM	-7.2	-0.1	6.2	-13.3	
dK	77.1	18.9	**28.7**	29.5	
dL	3.1	-9.3	28.7	-16.3	
dM	5.8	-9.6	28.7	-13.3	

table **10** of chapter 14

Single-factor Three-way Elementary Analysis of table 1
(Johnson and Tsao example)

	1	**4**	**7**	(x)	(x)	(·)
aK	-20.0	-16.8	-16.0	**-27.1**	**29.5**	**47.2**
a**L**	2.8	10.0	15.8	-27.1	**-16.3**	
a**M**	1.8	8.0	14.8	-27.1	**-13.3**	
bK	6.9	-11.9	-5.1	**-8.0**	29.5	
bL	-0.3	0.9	4.7	-8.0	-16.3	
bM	-4.3	1.9	7.7	-8.0	-13.3	
cK	8.7	2.9	-5.3	**6.2**	29.5	
cL	-3.5	-2.3	0.5	6.2	-16.3	
cM	-7.5	-0.3	7.5	6.2	-13.3	
dK	36.2	20.4	0.2	**28.7**	29.5	
dL	-6.0	-6.8	-15.0	28.7	-16.3	
dM	-14.0	-5.8	-9.0	28.7	-13.3	
(x)	**9.4**	**-1.8**	**-7.6**			

Examples:

-20.0 = -2.4 + (-17.6)
-16.8 = 0.8 + (-17.6)

2.8 = -6.7 + (9.5)

Term	**High**	**Low**	**Range**
[common]	[47.2]		
abcd	28.7	-27.1	56
KLM	29.5	-16.3	46
147	9.4	-7.6	17
Residual	36.2	-20.0	56

they both have to be included in the "residual" term. Our residuals in table 10 were thus obtained. When we need to be more precise, we shall call such residuals *after-single-factors residuals*, or even "residuals after a single-factor analysis".

Given the breakdown, the natural next step is to ask about rough sizes -- at least comparative sizes -- for the various terms. The simplest things to look at are highest and lowest values and the corresponding ranges. These are already given in the lower right corner of table 10.

While we have no doubt that the factors have substantial effects, it would be wrong to feel that, so far as the analysis of table 10 is concerned, these effects account for most of what is going on. Indeed, in these rough terms, the residual term is as large as the largest of the other terms.

14C. Moving onwards

In reaching our single-factor three-way elementary analysis, we had to take one backward step. The last column of table 9 -- consisting of that part of the adjustment for **abcd** and **KLM** jointly that was not accounted for by the separate **abcd** and **KLM** adjustments -- was found as something separate, and then thrown back into the after-single-factor residual. We need a name for terms of this sort -- joint adjustment less separate adjustments -- and will call them *interactions*, even, when we need to be precise, *two-factor interactions*.

In the trivial case of two-factor data, the two-factor interactions reduce to the usual residuals, residuals that are, more explicitly, after-single factor residuals.

We calculated an **abcd**x**KLM** interaction almost automatically, as the last part of an analysis that began with a compound **abcd**x**KLM** identification of rows. If we had rearranged the data, making **abcd**x**147** the compound identification of rows, the same arithmetic would have brought us an **abcd**x**147** interaction. Similarly for **KLM**x**147**.

Table **11** shows the after-single-factor residuals and the three two-factor interactions concealed in them. Notice that each two-factor interaction can also be found as the arithmetic mean of the corresponding after-single-factor residuals.

The way in which the two new interactions appear on the page is different from the way in which the first interaction appears. Each value in these lower two panels is associated with some -- not all -- of the values of the after-single-factor residuals in the corner panel which are directly above it. Thus -5.1, in row **a** and column **1** is associated with -20.0, 2.8, and 1.8 in rows **aK**, a**L**, a**M**, but not with the other nine residuals in column **1**. Similarly, -1.4, in row **K** and column **4**, is associated with -16.8, -11.9, 2.9, and 20.4 in rows **aK**, **b**K, **c**K, and **d**K, but not with the other eight residuals in column **4**. This linking by tag, where L goes with aL, bL, cL, dL, and M goes with aM, bM, cM, dM, etc. is an important element in making displays of three- and more-way analyses that we can come to comprehend.

Once the three two-factor interactions are calculated, it is easy to find the residuals -- naturally called a f t e r - t w o - f a c t o r r e s i d u a l s that will remain after they

table **11** of chapter 14

The After-single-factors Residuals of table 10 and the Three Two-factor Interactions Therein Concealed (Johnson and Tsao Example, Analyzed by Means)

		1	**4**	**7**	(xx)
	aK	-20.0	-16.8	-16.0	-17.6
	a**L**	2.8	10.0	15.8	9.5
	a**M**	1.8	8.0	14.8	8.2
	bK	6.9	-11.9	-5.1	-3.4
	bL	-0.3	0.9	4.7	1.8
	bM	-4.3	1.9	7.7	1.8
	cK	8.7	2.9	-5.3	2.1
	cL	-3.5	-2.3	0.5	-1.8
	cM	-7.5	-0.3	7.5	-0.1
	dK	36.2	20.4	0.2	18.9
	dL	-6.0	-6.8	-15.0	-9.3
	dM	-14.0	-5.8	-9.0	-9.6
	a	-5.1	0.4	4.9	
	b	0.8	-3.0	2.4	
(xx)	**c**	-0.8	0.1	0.9	
	d	5.4	2.6	-7.9	
	K	8.0	-1.4	-6.6	
(xx)	**L**	-1.8	0.6	1.5	
	M	-6.0	1.0	5.2	

Examples: $-5.1 = \frac{1}{3}(-20.0+2.8+1.8)$

$8.0 = \frac{1}{4}(-20.0+6.9+8.7+36.2)$

$-17.6 = \frac{1}{2}(-20.0-16.8-16.0)$

are removed. For the upper left element, for example, we have

$$- 20.0 - (-17.6) - (-5.1) - 8.0 = -5.3$$

continuing in this way, we find all the after-two-factors residuals -- all the residuals after all single-factor and two-factor terms have been removed. We now have all that is needed to fill out the full (three-way) two-factor analysis into eight terms that we promised in the introduction.

Table **12** provides this eight-term breakdown. Its structure follows the patterns of tables 10 and 11, but is complex and novel enough to deserve graphic explanation. Accordingly, table **13** displays two of the 36 sets of eight terms that are implicit in table 12, that provide a breakdown of each of the 36 given data values.

Turning back to table 12, we naturally begin by getting some appreciation of the relative sizes of the various terms. The simplest thing to do is just to note the largest and smallest values and their range for each term. This has already been done in the lower right-hand corner of table 12. Looking at this summary -- or most other summaries -- we easily conclude that there are four large terms and four moderate ones, the large ones being:

- common.

- **abcd** -- uniform adjustments for the four rates of increase.

- **KLM** -- uniform adjustments for the three observers.

- **abcd**x**KLM** -- the interaction between the two factors just mentioned.

table **12** of chapter 14

The Two-factor Three-way Elementary Analysis
of the Data of table 1
(Johnson and Tsao example, analyzed by means)

		1	**4**	**7**	(xx)	(x)	(x)	(·)
	aK	-5.3	1.8	3.3	-17.6	**-27.1**	**29.5**	47.2
	a**L**	0.2	-0.5	-0.1	9.5	-27.1	**-16.3**	
	a**M**	4.7	-1.6	-3.5	8.2	-27.1	**-13.3**	
	bK	1.5	-4.1	2.5	3.4	**-8.0**	29.5	
	bL	-1.1	1.5	-1.0	1.8	-8.0	-16.3	
	bM	-0.9	2.1	-1.7	1.8	-8.0	-13.3	
	cK	-0.6	2.1	-1.7	2.1	**6.2**	29.5	
	cL	0.9	-1.2	-0.1	-1.8	6.2	-16.3	
	cM	-0.6	-1.3	1.5	-0.1	6.2	-13.3	
	dK	3.9	0.3	-4.2	18.9	**28.7**	29.5	
	dL	-0.3	-0.7	0.7	-9.3	28.7	-16.3	
	dM	-3.8	0.2	3.3	-9.6	28.7	-13.3	
	a	-5.1	0.4	4.9				
(xx)	**b**	0.8	-3.0	2.4				
	c	-0.8	0.1	0.9				
	d	5.4	2.6	-7.9				
	K	8.0	-1.4	-6.6				
(xx)	**L**	-1.8	0.6	1.5				
	M	-6.0	1.0	5.2				
(x)		9.4	-1.8	-7.6				

Term	**High**	**Low**	**Range**
[common]	[47.2]		
abcd	28.7	-27.1	56
KLM	29.5	-16.3	46
147	9.4	-7.6	17
abcdx**KLM**	18.9	-17.6	36
abcdx**147**	5.4	-7.9	13
KLMx**147**	8.0	-6.6	15
residual*	4.7	-5.3	10

*after-two-factor residuals

table 13 of chapter 14

Two of the Sets of Eight Terms in table 12 That Combine to Give the Data Values

		1	4	7	(xx)	(x)	(x)	(·)
	aK	-5.3	.....		-17.6	..-27.1	... 29.5	..47.2
	aL	.						
	aM	.						
		.						
	bK	.						
	bL	.						
	bM	.						
		.						
	cK	.						
	cL	.						
	cM	.	-1.3		. -0.1	... 6.2 ..	-13.3	..47.2
		.	.					
	dK	.	.					
	dL	.	.					
	dM	.	.					
	a	-5.1	.					
(xx)	b	.	.					
	c	.	0.1					
	d	.	.					
	K	8.0	.					
(xx)	L	.	.					
	M	.	1.0					
(x)		9.4	-1.8					

	For aK1	For cM4
	-5.3	-1.3
	-5.1	0.1
	8.0	1.0
	9.4	**-1.8**
	-17.6	**-0.1**
	-27.1	6.2
	29.5	-13.3
	47.2	**47.2**
Sum	39	38

(These sums are the data values; see table 1)

Notice that, although four of the terms are moderate, none is really small.

14D. The ell and the block

There are various ways to rearrange the display of table 12. Some follow the same basic format, only changing the pattern of the original table -- always exactly reflected by the table of residuals -- on which all is hung. Others change the basic format.

As an example of the first, consider table **14**.

Tables 12 and 14 are two of the six equivalent forms -- or two of twelve if we want to also consider a mere interchange of rows and columns as making a difference. The six possibilities can be described as follows:

Columns	Rows, major	Rows, minor	**Note**
147	**abcd**	**KLM**	Table 12
147	**KLM**	**abcd**	
KLM	**147**	**abcd**	
KLM	**abcd**	**147**	Table 14
abcd	**KLM**	**147**	
abcd	**147**	**KLM**	

This type of choice we will always have with us, whatever other types of choice we may or may not have. In particular, these choices apply just as much to the single-factor analysis of table 10 as to the two-factor analysis of table 12.

table **14** of chapter 14

The Ell Display of the Two-factor Analysis on a Different Base

(Johnson and Tsao example analyzed by means)

		K	L	M	(xx)	(x)	(x)	(Ø)
	a1	-5.3	0.2	4.7	-5.1	**-27.1**	**9.4**	**47.2**
	a4	1.8	-0.5	-1.6	0.4	-27.1	**-1.8**	
	a7	3.3	-0.1	-3.5	4.9	-27.1	**-7.6**	
	b1	1.5	-1.1	-0.9	0.8	**-8.0**	9.4	
	b4	-4.1	1.5	2.1	-3.0	-8.0	-1.8	
	b7	2.5	-1.0	-1.7	2.4	-8.0	-7.6	
	c1	-0.6	0.9	**-0.6**	-0.8	**6.2**	9.4	
	c4	2.1	-1.2	-1.3	0.1	6.2	-1.8	
	c4	-1.7	-0.1	1.5	0.9	6.2	-7.6	
	d1	3.9	-0.3	-3.8	5.4	**28.7**	9.4	
	d4	0.3	-0.7	0.2	2.6	28.7	-1.8	
	d7	-4.2	0.7	3.3	-7.9	28.7	-7.6	
	a	-17.6	9.5	8.2				
(xx)	**b**	-3.4	1.8	1.8				
	c	2.1	-1.8	-0.1				
	d	18.9	-9.3	-9.6				
(xx)	**1**	8.0	-1.8	-6.0				
	4	-1.4	0.6	1.0				
	7	-6.6	1.5	5.2				
(x)		29.5	-16.3	-13.3				

Note that this is equivalent to table 12, overall and in detail. Time spent tracing how entries in one match entries in the other should provide better understanding of how such tables work.

When we are dealing with a full analysis -- for a three-way table this is a two-factor analysis; for a four-way table a three-factor analysis; and so on -- we can gain some compactness and a chance at different insights by a more basic change, by going from an ell structure to a block structure. By abandoning the ell structure, we lose a certain simplicity. It is no longer true that the individual values that combine to give a data value lie on an open L. Rather, they lie on a lattice -- for three-way analysis a 4x2 lattice. Both block and ell forms have their advantages, we ought to understand and be prepared to use both.

Table **15** shows the block form of table 12. Notice that the (xx), (x), (·) notations have had to be removed, since there are, in each row and column, panels of more than one type. The Ø notation, standing for "no tag", has continued with a different meaning. Instead of this type of tagging, we are able -- and need -- to go over to using bold entries in a way closely analogous to their use in elementary analysis of a two-way table. There we went light-bold-light for residual, stripe adjustments and common. Here we go light-bold-light-bold for residuals, two-factor terms, single-factor terms, and common. (The extension to more than three factors is obvious.)

Table **16** shows how the values into which individual data values have been broken are arranged. It has exactly the same relation to table 15 that table 13 has to table 12. Its study should give further insight into the block display of the results of a full analysis of a multi-way table.

* * *

table **15** of chapter 14

Block Form of the Two-factor (Full) Analysis
(Johnson and Tsao example analyzed by means)

	1	4	7	(Ø)
aK	-5.3	1.8	3.3	**-17.6**
aL	0.2	-0.5	-0.1	**9.5**
aM	4.7	-1.6	-3.5	**8.2**
bK	1.5	-4.1	2.5	**-3.4**
bL	-1.1	1.5	-1.0	**1.8**
bM	-0.9	2.1	-1.7	**1.8**
cK	-0.6	2.1	-1.7	**2.1**
cL	0.9	-1.2	-0.1	**-1.8**
cM	-0.6	-1.3	1.5	**-0.1**
dK	3.9	0.3	-4.2	**18.9**
dL	-0.3	-0.7	0.7	**-9.3**
dM	-3.8	0.2	3.3	**-9.6**
a	**-5.1**	**0.4**	**4.9**	-27.1
b	**0.8**	**-3.0**	**2.4**	-8.0
c	**-0.8**	**0.1**	**0.9**	6.2
d	**5.4**	**2.6**	**-7.9**	28.7
K	**8.0**	**-1.4**	**-6.6**	29.5
L	**-1.8**	**0.6**	**1.5**	-16.3
M	**-6.0**	**1.0**	**5.2**	-13.3
Ø	9.4	-1.8	-7.6	**47.2**

Ranges

	147	Ø
abcdxKLM	10	36
abcd	13	56
KLM	15	46
Ø	17	

Highs

4.7	18.9
5.4	28.7
8.0	29.5
9.4	-

Lows

-5.3	-17.6
-7.9	-27.1
-6.6	-16.3
-7.6	-

Full tags
Terms

Residual	abcdxKLM
abcdx147	abcd
KLMx147	KLM
147	*

*common = 47.2

table 16 of chapter 14

Two of the Sets of Eight Terms in table 15 That Combine to Give the Data Values

	1	4	7	.
aK	-5.3			**-17.6**
aL				
aM				
bK				
bL				
bM				
cK				
cL				
cM		-1.3		**-0.1**
dK				
dL				
dM				
a	**-5.1**			-27.1
b				
c		**0.1**		6.2
d				
K	**8.0**			29.5
L				
M		**1.0**		-13.3
.	9.4	1.8		**47.2**

-5.3	-17.6
-5.1	-27.1
8.0	29.5
9.4	47.2

sum to:
39

-1.3	-0.1
0.1	6.2
1.0	13.3
1.8	47.2

sum to:
38

Coming back to the block form of table 15, we ought also to take note how it reminds us that the various terms into which we are analyzing the data are of four different structures, namely:

- one term, the residuals, that requires a three-way table to state its values. (All the other terms CAN be stated as three-way tables, but can also be stated in more compressed form.)
- three terms, the two-factor interactions, that require two-way tables to state their values. (We must regard the long panel on the upper right of table 15 as a two-way table, although it is here given in strung-out form.)
- three terms, the single-factor terms, that require only one-way tables to state their values.
- one term, the common term, that requires only a no-way table -- a single number -- to state its values.

It will be worthwhile, as we go on to keep the distinction among three-way, two-way, one-way, and no-way terms as clearly in our mind as we can.

14E. Choice of expression

The largeness of one of the two-factor interactions makes it extremely clear that, as we had already suspected, the (three-way) single-factor analysis of table 10 is quite inadequate. We would be much happier if a single-factor analysis were adequate, since then:

- all the terms of the partial description -- terms that we ought to try to understand -- would be simple in form. Thus we could expect better understanding and hope for deeper.

- fewer terms would be needed to describe all the recognized behavior.

- the interactions would be available to trap specific behavior that, though not confined to a single value, was still localized so far as two factors were concerned.

To move toward a more thorough single-factor analysis would be very good. How can we try? We must always ask about the possible gain from a better choice of expression. Let us do this.

At the close of the last section, we identified **abcd**, **KLM**, and their interaction as the large terms. Presumably a better choice of expression would reduce this interaction. Let us, then, look at the two-way table involving these terms. Panel (b) of table 9 sets things out numerically. Table **17** shows the coded residuals. We do not have to look very hard at either to see the need -- and the opportunity -- for a better choice of expression.

If we want to go after this possibility on a full-dress basis, we want to find suitable plotting values and then plot our single-factor residuals against these plotting values. When we had two factors, we found (in chapter 6) that

table 17 of chapter 14

Coded Residuals for Panel (b) of table 9
(Johnson and Tsao example analyzed by means)

0	-27.1	-8.0	6.2	28.7	
29.5	LO·	···	···	HIG	K
-13.3	··H	···	···	L··	M
-16.3	··H	···	···	L··	L
	a	b	c	d	

$$\frac{(\text{row term})(\text{column term})}{(\text{constant term})}$$

was a good plotting value. If we try to think what would be natural for three factors, we will do well to come out with

$$\frac{(\textbf{abcd}\text{ term})(\text{KLM term}) + (\textbf{abcd}\text{ term})(\textbf{147}\text{ term}) + (\textbf{KLM}\text{ term})(\textbf{147}\text{ term})}{\text{constant term}}$$

which is, indeed, the appropriate plotting value to follow a three-way single-factor analysis.

In calculating such a plotting value, there is no need to be unduly exact in fixing the values of the terms that enter into the plotting value. Specifically we will do well enough, in the Johnson and Tsao example, by using the following approximations:

a	-30	**K**	30	**1**	10	**common**	50
b	-10	**L**	-15	**4**	0		
c	+10	**M**	-15	**7**	-10		
d	+30						

Table **18** shows the plotting values thus calculated. Plot **19** shows the after-single-factors residuals against the plotting values. The tightness of the points to a straight line is striking.

Since the slope indicated by this plot is very close to 1.0, we are urged (see section 6D) to try 1.0 lower power for y. Since 1 - 1.0 = 0.0, this means try log y. Table **20** shows the two-digit logs (using table 23 of chapter 2) and one form of an initial elementary analysis. To a hasty glance, the analysis seems to be going quite well here.

table **18** of chapter 14

Plotting Values and the Terms From Which
They Were Calculated (Common = 50)
(Johnson and Tsao modified)

	1	**4**	**7**	(x)	(x)
aK	-18	-18	-18	**-30**	**30**
a**L**	0	+9	+18	-30	**-15**
a**M**	0	+9	+18	-30	**-15**
bK	-2	-6	-10	**-10**	30
bL	-2	3	8	-10	-15
bM	-2	3	8	-10	-15
cK	14	6	-2	**10**	30
cL	-4	-3	-2	10	-15
cM	-4	-3	-2	10	-15
dK	30	18	6	**30**	30
dL	-6	-9	-12	30	-15
dM	-6	-9	-12	30	-15
(x)	10	0	-10		

Examples

$$-18 = \frac{(-30)(30) + (-30)(10) + (30)(10)}{50}$$

$$0 = \frac{(-30)(-15) + (-30)(10) + (-15)(10)}{50}$$

$$9 = \frac{(30)(-15) + (-30)(0) + (-15)(0)}{50}$$

plot **19** of chapter 14

After-single-factors residuals (see table 10) against plotting values (see table 18)

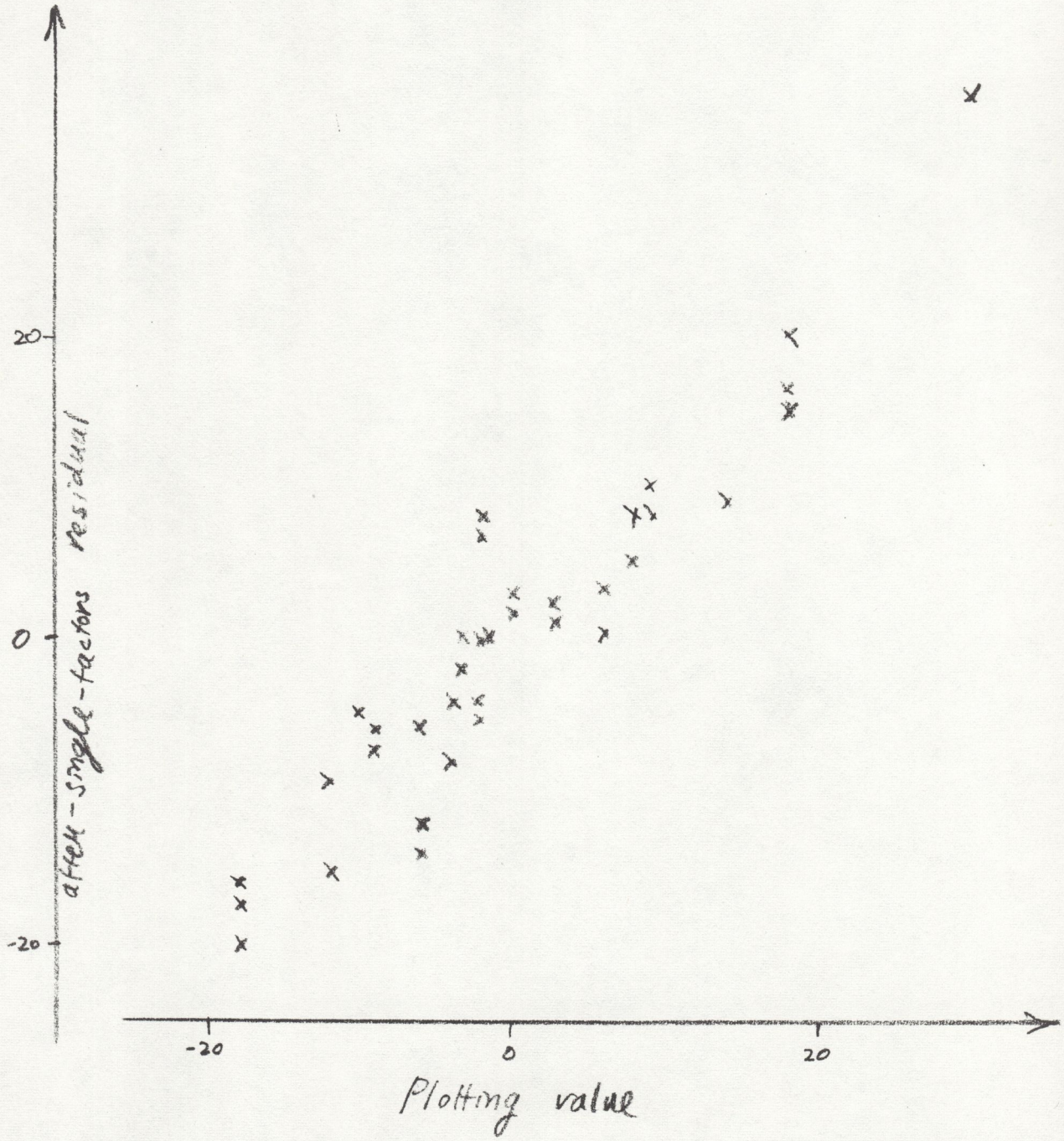

table 20 of chapter 14

Logarithmic Values of the Basic Observations (to left, entries are $-100 + 100 \log_{10}$) and the Results of One Form of Their Initial Elementary Analysis by Means (to the right)

Weight & Subject	Rate				Rate				
	a	b	c	d	a	b	c	d	
1K	59	93	100	118	-1.6	4.3	-3.2	0.5	34.2
4K	49	74	93	109	-0.4	-3.4	1.0	2.7	23.0
7K	42	75	85	99	-1.4	3.6	-1.0	-1.3	17.0
1L	20	51	63	80	-1.6	1.4	-1.2	1.5	-4.8
4L	8	34	52	71	-1.4	-3.4	0.0	4.7	-17.0
7L	8	30	48	57	4.1	-1.9	1.5	-3.8	-22.5
1M	26	49	62	76	4.6	-0.4	-2.0	-2.3	-5.0
4M	11	42	58	74	-3.4	-0.4	1.0	2.7	-12.0
7M	15	42	60	66	1.1	0.1	3.5	-4.8	-12.5
					-31.8	-3.8	10.8	25.1	58.2

When we go ahead and calculate the single-factor three-way elementary analysis, we reach table **21**. This is a much cleaner analysis than we reached in table 10. Instead of the range of the residuals equalling that of the largest single-factor term, their range is less than 1/4 as large. Where we were quite uncomfortable with a single-factor analysis for y, we can live reasonably well with a single factor analysis for log y.

Before digging into this partial description to see what we can learn from it, we may as well go on and write down the two-factor analysis. This is done in table **22**.

When we look at "high, low, and range", we see that we now have, omitting the common term,

- two large terms.
- one moderate-sized term.
- four small terms, consisting of three two-factor interactions and the after-two-factor residuals.

There are two directions in which we can naturally move, now that the more automatic part of the analysis is over and it is time to turn to interpretation:

- we can try to look at the two-factor analysis.
- we can return to the single-factor analysis and see what can be discovered there.

If we were to attack the two-factor analysis, we might do well to try to see the content of table 22 more clearly. One approach to this is to trim off the decimal points -- no

table **21** of chapter 14

(Three-way) Single-factor Analysis
(Johnson and Tsao example, expressed in logs)

	1	**4**	**7**	(x)	(x)	(·)
aK	-0.2	-0.1	-3.1	**-31.8**	**24.7**	58.2
a**L**	0.3	-1.6	2.4	-31.8	**-14.8**	
a**M**	1.3	-3.6	4.4	-31.8	**-9.8**	
bK	5.8	-3.1	1.9	**-3.8**	24.7	
bL	3.3	-3.6	-3.6	-3.8	-14.8	
bM	-3.7	-0.6	3.4	-3.8	-9.8	
cK	-1.8	1.3	-2.7	**10.8**	24.7	
cL	0.7	-0.2	-0.2	10.8	-14.8	
cM	-5.3	0.8	6.8	10.8	-9.8	
dK	1.9	3.0	-3.0	**25.1**	24.7	
dL	3.4	4.5	-5.5	25.1	-14.8	
dM	-5.6	2.5	-1.5	25.1	-9.8	
(x)	**8.1**	**-2.0**	**-6.0**			

Term	**High**	**Low**	**Range**
[common]	[58.2]		
abcd	25.1	-31.8	57
KLM	24.7	-14.8	40
147	8.1	-6.0	14
Residual	6.8	-5.6	12

table **22** of chapter 14

(Three-way)Two-factor Analysis

(Johnson and Tsao example, logs)

		1	**4**	**7**	(xx)	(x)	(x)	(·)
	aK	-0.9	2.5	-1.5	-1.2	**-31.8**	**24.7**	58.2
	a**L**	-2.4	0.0	2.5	0.3	-31.8	**-14.8**	
	a**M**	3.4	-2.4	-0.9	0.7	-31.8	**-9.8**	
	bK	1.1	-2.5	1.5	1.5	**-3.8**	24.7	
	bL	0.9	0.3	-1.2	-1.3	-3.8	-14.8	
	bM	-2.0	2.2	-0.3	-0.2	-3.8	-9.8	
	cK	-0.1	1.3	-1.3	-1.0	**10.8**	24.7	
	cL	0.8	-0.8	0.1	0.1	10.8	-14.8	
	cM	-0.8	-0.6	1.3	0.9	10.8	-9.8	
	dK	0.0	-1.3	1.4	0.6	**25.1**	24.7	
	dL	0.8	0.5	-1.3	0.8	25.1	-14.8	
	dM	-0.7	0.8	0.0	-1.5	25.1	-9.8	
	a	0.5	-1.7	1.3				
(xx)	**b**	1.8	-2.4	0.6				
	c	-2.1	0.7	1.3				
	d	-0.1	3.4	-3.3				
	K	1.4	1.9	-3.3				
(xx)	**L**	0.3	-0.2	-0.2				
	M	-1.7	-1.7	3.3				
		8.1	-2.0	-6.0				

Term	**High**	**Low**	**Range**
[common]	[58.2]		
abcd	25.1	-31.8	57
KLM	24.7	-14.8	40
147	8.1	-6.0	14
abcdx**KLM**	1.5	-1.5	3
abcdx**147**	3.4	-3.3	7
KLMx**147**	3.3	-3.3	7
residuals	3.4	-2.4	6

term is so little variable as to suggest that the decimals tell any key story. Table 23 shows the result of going a little further than this, dividing by two before rounding to integers. We leave it to the reader's tender mercies.

Turning back to the single-factor terms, we naturally try to make explicit just what the versions are, and how they relate. Table 24 sets out some of the facts.

Panel (a) treats the effect of rate of change of pull. A careful look shows us that the fitted term varies quite closely with the log of the rate, the differences "fitted minus log" being -31.8, -33.9, -36.9, and -35.1 with a mean of -34.4. Let us take this for an observed fact, and return to it in a moment.

Panel (b) deals with the three observers present in our limited set of data. Here it looks as though sex might be relevant -- though three cases are hardly enough to make any strong suggestion.

Panel (c) deals with the factor of initial pull, which is found to have some effect, though not a large one. Here the effect is crudely like $-\frac{1}{4}\log(\text{initial pull})$, the differences "fitted minus log" being 8.1, 8.0 and 9.0.

Thus some discussion, interesting and possibly useful can be made of each of the three single-factor terms.

In order to understand the first of these, it is time to note what we have been analyzing. In fact:

$$y = \text{change in pull before a change detected}$$

$$\log y = -100 + 100 \cdot \log_{10}(\text{change in grams})$$

table **23** of chapter 14

The Entries of table 22 Halved and Recorded to the Nearest Integer (+ = +1, - = -1)
(Johnson and Tsao Example, Logs)

		1	**4**	**7**	(xx)	(x)	(x)	(Ø)
	aK	0	+	-	-	**-16**	**12**	**29**
	a**L**	-	0	+	0	-16	**-7**	
	a**M**	+2	0	0	0	-16	**-5**	
	bK	0	-	+	+	**-2**	12	
	bL	0	0	-	-	-2	-7	
	bM	-	+	0	0	-2	-5	
	cK	0	+	-	0	**5**	12	
	cL	0	0	0	0	5	-7	
	cM	0	0	+	0	5	-5	
	dK	0	-	+	0	**13**	12	
	dL	0	0	-	0	13	-7	
	dM	0	0	0	-	13	-5	
	a	0	-	+				
(xx)	**b**	+	-	0				
	c	-	0	0				
	d	0	+2	-2				
	K	+	+	-2				
(xx)	**L**	0	0	0				
	M	-	-	+2				
(x)	•	+4	-	-3				

table **24** of chapter 14

The Single-factor Terms Made Explicit

a) THE **abcd** OR "RATE" FACTOR:

Version	**Rate** (gm/min)	**log rate***	**fitted term**
a	100	0.0	-31.8
b	200	30.1	-3.8
c	300	47.7	10.8
d	400	60.2	25.1

b) THE **KLM** OR "OBSERVER" FACTOR:

Version	**Sex**	**fitted term**
K	Male	24.7
L	Female	-14.8
M	Female	-9.8

c) THE **147** OR "INITIAL PULL" FACTOR:

Version	**Initial pull** (grams)	**log pull***	**fitted term**
1	100	0.0	8.1
4	250	39.8	-2.0
7	400	60.2	-6.0

---

* These columns are -200 + 100 $\log_{10}$(previous column).

Since the latter seems to move parallel to:

$$\text{log rate} = -200 + 100 \cdot \log_{10}(\text{grams per minute})$$

We ought to think about the -- little-changing -- difference:

$$\log y - \log \text{rate} = 100 + 100 \log_{10} \frac{\text{change in grams}}{\text{grams per minute}}$$

When we do this, we realize that

$$\frac{\text{change in grams}}{\text{grams per minute}} = \text{delay in minutes}$$

and that what we have observed can be very simply put in words as follows:

- the time taken to detect a beginning of a change in pull is much more constant than the change in pull taken to detect such a beginning.

Once we know this, there seems little reason not to take account of this from the beginning of our analysis. We ought to do better by analyzing

$$\log(\text{time of detection})$$

rather than either weight change for detection or log weight change. Doing this changes all initial entries by an amount that is uniform for each version of **abcd**. Since our fitting process is clearly transparent, this means that the residuals are unaffected, while the fit is changed in the exactly corresponding way.

Combining the common term and the rate term we have:

Version	Old fit	Change	New Fit
a	26.4	0.0	26.4
b	54.4	30.1	24.3
c	69.0	47.7	21.3
d	83.3	60.2	23.0

which leads to a common term of 23.8 and a much smaller rate term, whose values are 2.6, 0.5, -2.5, and -0.7. Tables **25** and **26** set out the corresponding single-factor and ell-form two-factor analyses.

When we ask what form of display is most likely to emphasize where we now stand, it is not unreasonable to notice that differences between observers are now the largest thing in sight. We might like, then, to choose a display that will make it easy to separate person-to-person differences -- both single-factor and interaction -- from those terms which do not explicitly involve observers (because we have averaged observers out). Such a display is easily had. We need only choose the most useful block form -- one in which observers are the whole of either rows or columns, say columns -- and then split the block between person vs person differences (all kinds) on the one hand, and all else (averaged over persons) on the other.

Table **27** shows the result. At first glance, we see a relatively clean analysis, however we choose to look:

- overall.
- at the person-vs-person column only.
- at the all-else column only.

table 25 of chapter 14

Single-factor Analysis

(Johnson and Tsao data, log of delay)

	1	4	7	(x)	(x)	(·)
aK	-0.2	-0.1	-3.1	2.6	24.7	23.8
aL	0.3	-1.6	2.4	2.6	-14.8	
aM	1.3	-3.6	4.4	2.6	-9.8	
bK	5.8	-3.1	1.9	0.5	24.7	
bL	3.3	-3.6	-3.6	0.5	-14.8	
bM	-3.7	-0.6	3.4	0.5	-9.8	
cK	-1.8	1.3	-2.7	-2.5	24.7	
cL	0.7	-0.2	-0.2	-2.5	-14.8	
cM	-5.3	0.8	6.8	-2.5	-9.8	
dK	1.9	3.0	-3.0	-0.7	24.7	
dL	3.4	4.5	-5.5	-0.7	-14.8	
dM	-5.6	2.5	-1.5	-0.7	-9.8	
(x)	8.1	-2.0	-6.0			

table **26** of chapter 14

Two-factor Analysis -- Ell-form
(Johnson and Tsao Data, Log of Delay)

		1	**4**	**7**	(xx)	(x)	(x)	(Ø)
	aK	-0.9	2.5	-1.5	-1.2	**2.6**	**24.7**	23.8
	a**L**	-2.4	0.0	2.5	0.3	2.6	**-14.8**	
	a**M**	3.4	-2.4	-0.9	0.7	2.6	**-9.8**	
	bK	1.1	-2.5	1.5	1.5	**0.5**	24.7	
	bL	0.9	0.3	-1.2	-1.3	0.5	-14.8	
	bM	-2.0	2.2	-0.3	-0.2	0.5	-9.8	
	cK	-0.1	1.3	-1.3	-1.0	**-2.5**	24.7	
	cL	0.8	-0.8	0.1	0.1	-2.5	-14.8	
	cM	-0.8	-0.6	1.3	0.9	-2.5	-9.8	
	dK	0.0	-1.3	1.4	0.6	**-0.7**	24.7	
	dL	0.8	0.5	-1.3	0.8	-0.7	-14.8	
	dM	-0.7	0.8	0.0	-1.5	-0.7	-9.8	
	a	0.5	-1.7	1.3				
	b	1.8	-2.4	0.6				
(xx)	**c**	-2.1	0.7	1.3				
	d	-0.1	3.4	-3.3				
	K	1.4	1.9	-3.3				
(xx)	**L**	0.3	-0.2	-0.2				
	M	-1.7	-1.7	3.3				
(x)		8.1	-2.0	-6.0				

Term	**High**	**Low**	**Range**
[common]	[23.8]		
KLM	24.7	-14.8	40
147	8.1	-6.0	14
abcd	2.6	-2.5	5
KLMx147	3.3	-3.3	7
KLMxabcd	1.5	-1.5	3
147xabcd	3.4	-3.3	7
Residual	3.4	-2.5	6

table 27 of chapter 14

Two-factor Analysis -- Split Block Form, Person Centered
(Johnson and Tsao Data, Log of Delay)

	PERSON VS PERSON				ALL ELSE
	K	L	M	...	Ø
a1	-0.9	-2.4	3.4		0.5
a4	2.5	0.0	-2.4		-1.7
a7	-1.5	2.5	-0.9		1.3
b1	1.1	0.9	-2.0		1.8
b4	-2.5	0.3	2.2		-2.4
b7	1.5	-1.2	-0.3		0.6
c1	-0.1	0.8	-0.8		-2.1
c4	1.3	-0.8	-0.6		0.7
c7	-1.3	0.1	1.3		1.3
d1	0.0	0.8	-0.7		-0.1
d4	-1.3	0.5	0.8		3.4
d7	1.4	-1.3	0.0		-3.3
				...	
a	-1.2	0.3	0.7		2.6
b	1.5	-1.3	-0.2		0.5
c	-1.0	0.1	0.9		-2.5
d	0.6	0.1	-1.5		-0.8
				...	
1	1.4	0.3	-1.7		8.1
4	1.9	-0.2	-1.7		-2.0
7	-3.3	-0.2	3.3	...	-6.0
Ø	24.7	-14.8	-9.8		23.8
				...	

Ranges

	KLM	Ø
abcdx147	5.9	6.7
abcd	3.0	5.1
147	6.6	14.1
Ø	39.5	[23.8]

When we look in more detail, we notice that the largest interaction involving people KLMx147 does show a structure mildly suggestive of a desire for a different expression. Blithe calculation of plotting values and construction of a graph would lead to trying something like the square root of the present expression. Since this would mean using

$$\sqrt{\text{log delay}}$$

and since square roots of logarithms seem to be rather rarely useful, the appropriate response, which we leave to the reader, would seem to be:

- go back to the beginning, convert individual grams of change before detection to times of delay.
- reanalyze, exploring various alternative expressions, such as $-1/\sqrt{\text{delay}}$, as well as log delay.

The process of stripping off layer after layer is still going on.

14F. Inconsistent factors

One of the important properties of the three factors in the Johnson and Tsao example is that they are consistent -- in the sense that each of their versions has a consistent meaning wherever it appears in the data.

As many examples show, inconsistent factors are common. When dealing with them, we dare not make a full three-way analysis. Indeed, we may not even be able to separate out

all the single-factor terms. Particularly when dealing with more than three factors, a wide variety of detailed situations are possible. For three factors, only two are common, and examples of these may help to clarify matters.

* * *

An experiment studied the homogeneity of copper content among bronze castings made from the same pour. Two samples (A and B) were taken from each of 11 castings, and each sample was analyzed in duplicate. (Wernimont 1947, Bennett and Franklin, pp. 358-359). The results are shown in table 28. What analysis is it reasonable to make?

Naive following of marks on paper could lead to analysis as a three-way table: Castings 1 to 11, Samples A and B, Determinations 1 and 2. But a little thought suggests that we have no ground for such an analysis. So far as we know, Sample A from one casting is no more like Sample A from another than it is like Sample B from the same other. To use A-vs-B as a consistent difference is thus probably an error.

Something similar can be said about determinations. There need be no reason why determination 1 on Sample A should be more like determination 1 or more like determination 2 even on Sample B from the same casting. (We will return to what may actually be the case shortly.) Let us suppose that this is so.

Which terms of the conventional three-way analysis is it appropriate to fit? Consider first the two-way table in which castings fix rows, and samples and analyses combine to

table 28 of chapter 14

Copper Content of Bronze Castings

(excess over 85% in units of 0.01%)

Casting & Sample	Analysis 1	Analysis 2
1A	54	56
1B	51	54
2A	54	60
2B	25	25
3A	72	77
3B	-6	-5
4A	48	50
4B	-2	2
5A	54	57
5B	84	84
6A	72	86
6B	81	91
7A	72	76
7B	81	84
8A	12	12
8B	12	20
9A	47	49
9B	75	77
10A	-2	10
10B	90	90
11A	12	17
11B	18	24

Source: G. Wernimont 1947, quoted by C. A. Bennett and N. L. Franklin 195? Statistical Analysis in Chemistry and Chemical Industry, pp. 358-359. John Wiley and Sons.

fix columns. We cannot fit column means, for there is nothing consistent about, say, the combination of sample A and analysis 2. We can do an incomplete elementary analysis, taking matters as far as panel **a** in table **29** suggests. This will give us:

- a common term.
- a casting term.
- remainder combining sample and analysis differences.

Within the remainder, we pair the two analyses associated with a single sample. If we fit a mean to each such pair, we can separate out a term to describe these means, reaching the analysis shown in panel **b** of table 29. Sober consideration shows that this is as far as we can go with this example.

* * *

Another experiment studied the values found for the "total solid content" of wet brewers' yeast when the technique of measurement was changed. Specifically, the length of a drying period was varied. (Bennett and Franklin, pp. 363-369). Presumably we anticipate the longer drying will remove more water and other volatile materials, and thus lead to somewhat smaller residues and, accordingly, somewhat smaller "percent total solids."

In this experiment, each of 10 different samples of the wet brewers' yeast was dried, in duplicate, for each of three durations. The data is shown in table **30**. What analysis does it deserve?

table 29 of chapter 14

Stages in the Analysis of the Bronze Casting Example
(Four-term Break)

a) OUTLINE OF INCOMPLETE ELEMENTARY ANALYSIS

Casting	A1	A2	B1	B2	(x)
1					
2					
...					
11					
					(·)

b) RESULT OF FINDING AND SEPARATING MEANS OF TWO DETERMINATIONS ON EACH SAMPLE

	A1	A2	B1	B2	A	B	(x)	(·)
1	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-
...								
11	-	-	-	-	-	-	-	

c) NAMES FOR THE BREAK OF PANEL b

Residuals	Sample x Casting	Casting	Common

table 30 of chapter 14

Percent Total Solids Based on Different Durations of Drying

Sample*	Duration of drying period		
	3 hours	6 hours	9 hours
1a	13.24	13.16	12.96
1b	13.56	13.26	13.01
2a	13.92	13.81	13.76
2b	13.86	13.80	13.75
3a	19.13	18.86	18.70
3b	19.23	18.79	18.75
4a	18.35	18.11	17.94
4b	18.29	18.24	17.99
5a	15.51	15.06	14.84
5b	15.53	15.11	14.80
6a	16.63	16.61	16.60
6b	16.65	16.57	16.55
7a	19.29	18.96	18.84
7b	19.28	19.12	19.03
8a	18.76	18.39	18.23
8b	18.72	18.43	18.27
9a	18.03	17.86	17.72
9b	18.11	17.84	17.79
10a	16.61	16.32	16.21
10b	16.77	16.23	16.13

* a and b identify replications.

Samples and duration are surely factors of general applicability, with a common meaning. Duplicates are probably not such (unless a and b should happen to reflect something like order of test or which one of two ovens).

If samples fix columns, while duration and duplicates combine to fix rows, a complete two-way analysis is appropriate, giving us:

- a common term.
- a sample term.
- a joint duration-duplicate term.
- a remainder.

The natural next step is to write the duration-duplicate term as a two-way table. Once we do this, we need to remind ourselves that "duplicate" is an inconsistent factor, and that only a partial analysis is reasonable.

Similarly we can extract a duration x sample interaction term from the remainder, but dare not try to fit a duplicate x sample term. The shape of the resulting analysis is thus as in table **31**.

Should we actually want to make this analysis, and should we be wholly uninterested in the sample term, we could adjust the values for each sample separately without affecting the rest of the analysis. In particular we could work with the values of table **32**. Even a general look at this table gives a rather clear picture of what is actually going on.

table 31 of chapter 14

Outline of Analysis for Drying Duration Data
(Six-term Break)

a) THE OUTLINE:

		1	2	...	10	(xx)	(x)	(·)
	3a	-	-	...	-	-	-	-
	6a	-	-		-	-	-	|
	9a	-	-		-	-	-	|
	3b	-	-	...	-	-	-	|
	6b	-	-		-	-	-	|
	9b	-	-		-	-	-	|
	3	-	-	...	-			
(xx)	6	-	-		-			
	9	-	-					
(x)								

b) THE NAMES:

Residual:	Duplicates within*	Samples	Common
Duration x Samples			
Samples			

* Duplicates within = duplicates within combinations samples and durations.

table 32 of chapter 14

Values of table 30 Adjusted for Easy Calculation
(Sample-to-sample differences modified. All values in 0.01%)

(Adjustment*)	Sample	Duration 3 hours	6 hours	9 hours
(13.20)	1a	4	-4	-24
	1b	36	6	-19
(13.80)	2a	12	1	-4
	2b	6	0	-5
(18.80)	3a	33	6	-10
	3b	43	-1	5
(18.20)	4a	15	-9	-24
	4b	9	4	-21
(15.10)	5a	41	-4	-26
	5b	43	1	-30
(16.60)	6a	3	1	0
	6b	5	-3	-5
(19.00)	7a	29	-4	-16
	7b	28	12	3
(18.40)	8a	36	-1	-17
	8b	32	3	-13
(17.80)	9a	23	-4	-8
	9b	31	-6	-1
(16.30)	10a	31	2	-9
	10b	47	-7	-17

14G. Sweeping out with elementary analyses

Our description of how to do a three-way analysis has a strong flavor of catch-as-catch-can, of begin where you like and keep on until you have what you need. For this there is good reason: a single-factor analysis may be good enough. If it is, a catch-as-catch-can approach is both easy and natural.

If we plan to make a full analysis, whether of a three-way table or a more-way table, however, we can think about the process somewhat more effectively if we think of setting about our analysis in a more systematic way.

For a running start, let us go back to the analysis of two-way -- and even one-way tables. These analyses can, if we wish, be thought of as a sweeping out process in which we sweep as much into as few-way terms as we can. The one-way analysis

data = common + residual

is an analysis into:

- a no-way term, the typical value common to every data value.
- a one-way term, the residuals.

Good analysis is doing well in sweeping; sweeping the right part of the data into the no-way term, leaving in the residuals -- as well as we can -- only what belongs there.

For a two-way analysis

data = common + row adj. + col adj. + residual

we have:

- one no-way term, the common.
- two one-way terms, the stripe adjustments.
- one two-way term, the residuals.

The first step in a good analysis is to sweep the right amount out of data into one-way terms -- these can be temporary one-way terms combining part of what will be common with what will be stripe adjustment. Once we have done this, it is a more leisurely task to sweep out of our temporary one-way terms what should be included in the common term.

For a three-way analysis, we have:

- one no-way term, the common.
- three one-way terms, the single-factor terms.
- three two-way terms, the two-factor interactions.
- one three-way terms, the residuals.

Our first step can be to sweep out of the data -- into temporary two-way (and smaller) terms -- all that ought to be swept, as well as we can see it. Then we can sweep out of each of these two-way terms -- into temporary one- and no-way terms -- all that ought to be swept, as well as we can see it. Then we can repeat this again with the one-way terms.

All these sweeping processes, except the sweeping of one-way tables, can be carried out by making elementary analyses on a suitably chosen (often compound) two-way table.

* * *

Let us go through this procedure for the logarithmic form of the Johnson and Tsao example, keeping all stages of our computation in block form. Table **33** shows the data, while tables **34** to **41** show the successive stages of the procedure. Sweeping out by elementary analysis takes place:

- once at stage 2.
- once at stage 4.
- once at stage 6.
- twice at stage 7.
- once at stage 9 (not needed in this instance).

Except for arithmetic labor, nothing keeps us from going through this same procedure, using midmean differences in each elementary analysis.

Presumably, if we did this, we would want to use midmeans in sweeping pieces of the common term out of the one-way terms. As a result, both of the use of midmean differences for two-way terms and of this, we would find most of the 0.0's that appear in tables 36 to 41 replaced by numbers not exactly zero -- something worthy of note but not troublesome.

Even with a so small a three-way table as 3x3x4 = 36, using midmean differences takes us beyond the scope of comfortable hand arithmetic, since we need to handle two 3x12 analyses and one 4x9.

Actually, the anslysis by means is surprisingly clean. Table **42** gives the q-typical squares and typical midsquares.

* * *

table 33 of chapter 14

Stage 1: The Data

	a	b	c	d	.
1K	59	93	100	118	
4K	49	74	93	109	
7K	42	75	85	99	
1L	20	51	63	80	
4L	8	34	52	71	
7L	8	30	48	57	
1M	26	49	62	76	
4M	11	42	58	74	
7M	15	42	60	66	
1					
4					
7					
K					
L					
M					
.					

table 34 of chapter 14

Stage 2: One Elementary Analysis

	a	b	c	d	.
1K	-1.6	4.3	-3.2	0.5	34.2
4K	-0.4	-3.4	1.0	2.7	23.0
7K	-1.4	3.0	-1.0	-1.3	17.0
1L	-1.6	1.4	-1.2	1.5	-4.8
4L	-1.4	-3.4	0.0	4.7	-17.0
7L	4.1	-1.9	1.5	-3.8	-22.5
1M	4.6	-0.4	-2.0	-2.3	-5.0
4M	-3.4	-0.4	1.0	2.7	-12.0
7M	1.1	0.1	3.5	-4.8	-12.5
1					
4					
7					
K					
L					
M					
.	-31.8	-3.8	10.8	25.1	58.2

table 35 of chapter 14

Stage 3: Rearrangement of table 34

	1	4	7	.
aK	-1.6	-0.4	-1.4	
aL	-1.6	-1.4	4.1	
aM	4.6	-3.4	1.1	
bK	4.3	-3.4	3.6	
bL	1.4	-3.4	-1.9	
bM	-0.4	-0.4	0.1	
cK	-3.2	1.0	-1.0	
cL	-1.2	0.0	1.5	
cM	-2.0	1.0	3.5	
dK	0.5	2.7	-1.3	
dL	1.5	4.7	-3.8	
dM	-2.3	2.7	-4.8	
a				-31.8
b				-3.8
c				10.8
d				25.1
K	34.2	23.0	17.0	
L	-4.8	-17.0	-22.5	
M	-5.0	-12.0	-12.5	
.				58.2

table 36 of chapter 14

Stage 4: Another Elementary Analysis
(new sweepings bold)

	1	4	7	.
aK	-0.5	0.7	-0.3	**-1.1**
aL	-2.0	-1.8	3.7	**0.4**
aM	3.8	-4.4	0.3	**0.8**
bK	2.8	-4.9	2.1	**1.5**
bL	2.7	-2.1	-0.6	**-1.3**
bM	-0.2	-0.2	0.3	**-0.2**
cK	-2.1	2.1	0.1	**-1.1**
cL	-1.3	-0.1	1.4	**0.1**
cM	-2.8	0.2	2.7	**0.8**
dK	-0.1	2.1	-1.9	**0.6**
dL	0.7	3.9	-4.6	**0.8**
dM	-0.8	4.2	-3.3	**-1.5**
a				-31.8
b				-3.8
c				10.8
d				25.1
K	34.2	23.0	17.0	
L	-4.8	-17.0	-22.5	
M	-5.0	-12.0	-12.5	
	0.0	**0.0**	**0.0**	58.2 + **0.0**

table 37 of chapter 14

Stage 5: Another Rearrangement

	K	L	M	.
a1	-0.5	-2.0	3.8	
a4	0.7	-1.8	-4.4	
a7	-0.3	3.7	0.3	
b1	2.8	2.7	-0.2	
b4	-4.9	-2.1	-0.2	
b7	2.1	-0.6	0.3	
c1	-2.1	-1.3	-2.8	
c4	2.1	-0.1	0.2	
c7	0.1	1.4	2.7	
d1	-0.1	0.7	-0.8	
d4	2.1	3.9	4.2	
d7	-1.9	-4.6	-3.3	
a	-1.1	0.4	0.8	-31.8
b	1.5	-1.3	-0.2	-3.8
c	-1.1	0.1	0.8	-10.8
d	0.6	0.8	-1.5	25.1
1	34.2	-4.8	-5.0	0.0
4	23.0	-17.0	-12.0	0.0
7	17.0	-22.5	-12.5	0.0
.				58.2

table 38 of chapter 14

Stage 6: Another Elementary Analysis of the Upper Left Panel (new sweepings bold)

	K	L	M	∘
a1	-0.9	-2.4	3.4	**0.4**
a4	2.5	0.0	-2.6	**-1.8**
a7	-1.5	2.5	-0.9	**1.2**
b1	1.0	0.9	-2.0	**1.8**
b4	-2.5	0.3	2.2	**-2.4**
b7	1.5	-1.2	-0.3	**0.6**
c1	0.0	0.8	-0.7	**-2.1**
c4	1.4	-0.8	-0.5	**0.7**
c7	-1.3	0.0	1.3	**1.4**
d1	0.0	0.8	-0.7	**-0.1**
d4	-1.3	0.5	0.8	**3.4**
d7	1.4	-1.3	0.0	**-3.3**
a	-1.1	0.4	0.8	-31.8
b	1.5	-1.3	-0.2	-3.8
c	-1.1	0.1	0.8	10.8
d	0.6	0.8	-1.5	25.1
1	34.2	-4.8	-5.0	0.0
4	23.0	-17.0	-12.0	0.0
7	17.0	-22.5	-12.5	0.0
∘	**0.0**	**0.0**	**0.0**	58.2 + **0.0**

table 39 of chapter 14

Stage 7: Elementary Analyses of Two Two-way Tables in the Lower Left Corner

	K	L	M	c
a1	-0.9	-2.4	3.4	0.4
a4	2.5	0.0	-2.6	-1.8
a7	-1.5	2.5	-0.9	1.2
b1	1.0	0.9	-2.0	1.8
b4	-2.5	0.3	2.2	-2.4
b7	1.5	-1.2	-0.3	0.6
c1	0.0	0.8	-0.7	-2.1
c4	1.4	-0.8	-0.5	0.7
c7	-1.3	0.0	1.3	1.4
d1	0.0	0.8	-0.7	-0.1
d4	-1.3	0.5	0.8	3.4
d7	1.4	-1.3	0.0	-3.3
a	-1.1	0.4	0.8	-31.8 + 0.0
b	1.5	-1.3	-0.2	-3.8 + 0.0
c	-1.0	0.2	0.9	10.8 - 0.1
d	0.6	0.8	-1.5	25.1 + 0.0
1	1.4	1.9	-3.3	0.0 + 8.1
4	0.3	-0.2	-0.2	0.0 - 2.0
7	-1.7	-1.5	3.3	0.0 - 6.0
.	0.0	0.0	0.0	58.2
	+0.0	+0.0	+0.0	+0.0
	+24.7	-14.8	-9.8	+0.0

table 40 of chapter 14

Stage 8: Rearrangement Yet Again

	1	4	7	Ø
aK	-0.9	2.5	-1.5	-1.1
aL	-2.4	0.0	2.5	0.4
aM	3.4	-2.6	-0.9	0.8
bK	1.0	-2.5	1.5	1.5
bL	0.9	0.3	-1.2	-1.3
bM	-2.0	2.2	-0.3	-0.2
cK	0.0	1.4	-1.3	-1.0
cL	0.8	-0.8	0.0	0.2
cM	-0.7	-0.5	1.3	0.9
dK	0.0	-1.3	1.4	0.6
dL	0.8	0.5	-1.3	0.8
dM	-0.7	0.8	0.0	-1.5
a	0.4	-1.8	1.2	-31.8
b	1.8	-2.4	0.6	-3.8
c	-2.1	0.7	1.4	10.7
d	-0.1	3.4	-3.3	25.1
K	1.4	0.3	-1.7	24.7
L	1.9	-0.2	-1.5	-14.8
M	-3.3	-0.2	3.3	-9.8
Ø	8.1	-2.0	-6.0	58.2

table **41** of chapter 14

Stage 9: Elementary Analysis of **abcd** x **147**

(No changes required in table 40)

Stage 10: Sweeping out of one-way terms

(No changes required in table 40)

table **42** of chapter 14

q-typical Squares and Typical Midsquares
for the Analysis of tables 40-41

Term	**q-typical square**	**Typical midsquare**
abcd	380.	540.
KLM	200.	240.
147	40.	33.
abcd x **KLM**	1.7	.99
abcd x **147**	5.1	3.0
KLM x **147**	5.1	4.0
residuals	2.8	2.4
Single-factor residuals (sum of terms above)	15	10.
Single-factor residuals (table 21)	15	11

Entries above to leading two figures only. Actual values include: 377.8, 539.6, 198.3, 238.0, 39.76, 33.13, 1.66, .99, 5.14, 3.04, 5.12, 3.96, 2.85, 2.35.

Going on to four-way -- or nine-way -- tables needs no new insights, especially if we are prepared to think in terms of "sweeping out". For a full analysis of a four-way table, for instance, we have to deal with:

- 1 no-way term, the common term.
- 4 one-way terms, single-factor terms.
- 6 two-way terms, two-factor interactions.
- 4 three-way terms, three-factor interactions.
- 1 four-way term, the residuals.

We make the four-way term by sweeping -- or, if you prefer, combing, temporary three-way (and smaller) terms out of it. Then we make the final three-way terms by sweeping temporary two-way (and smaller) terms out of them, and so on.

More ways, more stages, much more arithmetic, nothing new.

14H. <u>How far have we come</u>?

This chapter offers the essential ideas that are needed to take apart -- and understand -- multiway tables. There are really two such:

- to the extent that factors are consistent, a complete breakdown involves one term for each set of one or more factors. What is described by a specific term is limited to what is not described by simpler terms involving some but not all of the same factors.

- however many factors there may be, we can always, if we have a large enough sheet of paper, write the data in row by column form, using compound tags.

Equally, we can write out the terms of the analysis -- in either ell or block form -- in a pattern fixed by the pattern chosen for the data -- again if we have a big enough sheet of paper. (This time it has to be bigger.)

Though these ideas are vital and dominant, all the familiar considerations are with us. We want to describe as much as we can using as little as we can -- choosing the expression of the data wisely still helps. Resistance to erratic values is still vital, though we have given no explicit attention to how to get it in the part of the chapter under review. The usual process of data analysis, stripping off layer after layer, continues to be prominent. Getting the answers into a form where we can look at -- and think about -- them continues to be vital.

14K. <u>One way to add resistance</u>

We have noticed that

- all our examples have used analyses by means.
- yet it is desirable to make resistant analyses.
- if we can face the arithmetic -- as when a computer will do it for us -- using midmean differences throughout is satisfactory.

Our purpose here is to illustrate a simple procedure that can often be used by hand. Its main disadvantages are:

- absence of symmetry, so that the result depends, somewhat, on the order in which the procedure is followed.
- a lesser degree of resistance than a full analysis by midmean differences.

Since the procedure is a linear descendant of the "streamlined" procedure for two-way tables, it will be convenient to call it, too, a "streamlined" procedure.

The basic idea is simple:

- first a full analysis by means.
- then a modification of this analysis, using midmeans, once in each possible direction.

Let us start with the mean analysis of tables 40-41, and begin by using midmeans "in the **147** direction" to readjust the analysis wherever possible. Thus in the top line of table 40, we see

1	**4**	**7**	**ø**
-0.9	2.5	-1.5	-1.1

which we readjust to

0.0	3.4	-0.6	-2.0

The new figures combine to give the same sums

-2.0	1.4	-2.6

as the old, but their breakdown has been altered to make the midmean of the three distinguish as **1**, **4**, and **7** vanish. We do a similar thing in every line of table 40, ending up with the conversion of

1	4	7	Ø
8.1	-2.0	-6.0	58.2

to

10.1	0.0	-4.0	56.2

Table 43 shows the result.

We now turn to the **KLM** direction. In the first column we have five sets of values involving **K, L, M,** and nothing or **Ø** together with other symbols, namely

aK	0.0	**bK**	0.0	**cK**	0.0	**dK**	0.0	**K**	1.1
aL	-2.4	**bL**	0.6	**cL**	0.8	**dL**	0.3	**L**	2.1
aM	4.3	**bM**	-1.7	**cM**	-0.2	**dM**	-0.7	**M**	-3.1
a	0.0	**b**	1.2	**c**	-2.8	**d**	0.0	**Ø**	10.1

Of these, only the last requires adjustment to bring the midmean to zero. Making all such readjustments wherever needed in any column, gives us the results of table 44.

We are now ready to repeat the process one more time, sweeping out midmeans in the **abcd** direction. In the first column, we now look at examples of the pattern **a, b, c, d** and nothing or **Ø**, of which there are the following four

aK	0.0	**aL**	-2.4	**aM**	4.3	**a**	0.0
bK	0.0	**bL**	0.6	**bM**	-1.7	**b**	1.2
cK	0.0	**cL**	0.8	**cM**	-0.2	**c**	-2.8
dK	0.0	**dL**	0.3	**dM**	-0.7	**d**	0.0
K	0.0	**L**	1.0	**M**	-4.2	**Ø**	11.2

of which the second and third require readjustment -- by 0.4 and -0.4 respectively. Making all such adjustments in all columns yields the entries of table 45.

table 43 of chapter 14

Results of Sweeping Out Midmeans from table 40
in the 147 Direction

	1	4	7	ø
aK	0.0	3.4	-0.6	-2.0
aL	-2.4	0.0	2.5	0.4
aM	4.3	-1.7	0.0	-0.1
bK	0.0	-3.5	0.5	2.5
bL	0.6	0.0	-1.2	-1.0
bM	-1.7	2.5	0.0	-0.5
cK	0.0	1.4	-1.3	-1.0
cL	0.8	-0.8	0.0	0.2
cM	-0.2	0.0	1.8	0.4
dK	0.0	-1.3	1.4	0.6
dL	0.3	0.0	-1.8	1.3
dM	-0.7	0.8	0.0	-1.5
a	0.0	-2.2	0.8	-31.4
b	1.2	-3.0	0.0	-3.2
c	-2.8	0.0	0.7	11.4
d	0.0	3.5	-3.2	25.0
K	1.1	0.0	-2.0	25.0
L	2.1	0.0	-1.3	-15.0
M	-3.1	0.0	3.5	-10.0
ø	10.1	0.0	-4.0	56.2

table **44** of chapter 14

Results of Sweeping Out Midmeans from table 43
in the **KLM** Direction

	1	**4**	**7**	**ø**
aK	0.0	3.4	-0.6	-1.9
aL	-2.4	0.0	2.5	0.5
aM	4.3	-1.7	0.0	0.0
bK	0.0	-3.5	0.5	3.0
bL	0.6	0.0	-1.2	-0.5
bM	-1.7	2.5	0.0	0.0
cK	0.0	1.4	-1.3	-1.2
cL	0.8	-0.8	0.0	0.0
cM	-0.2	0.0	1.8	0.2
dK	0.0	-1.3	1.4	0.0
dL	0.3	0.0	-1.8	0.7
dM	-0.7	0.8	0.0	-2.1
a	0.0	-2.2	0.8	-31.5
b	1.2	-3.0	0.0	-3.7
c	-2.8	0.0	0.7	11.6
d	0.0	3.5	-3.2	25.6
K	0.0	0.0	-0.7	35.0
L	1.0	0.0	0.0	-5.0
M	-4.2	0.0	4.8	0.0
ø	11.2	0.0	-5.3	46.2

table **45** of chapter 14

Results of Sweeping Out Midmeans from table 44

in the **abcd** Direction

	1	**4**	**7**	**ø**
aK	0.0	3.4	-0.6	-1.3
aL	-2.8	0.0	3.1	0.3
aM	4.7	-2.1	0.0	0.0
bK	0.0	-3.5	0.5	3.6
bL	0.2	0.0	-0.6	-0.7
bM	-1.3	2.1	0.0	0.0
cK	0.0	1.4	-1.3	-0.6
cL	0.4	-0.8	+0.6	-0.2
cM	0.2	-0.4	1.8	0.2
dK	0.0	-1.3	1.4	0.6
dL	-0.1	0.0	-1.2	0.5
dM	-0.3	0.4	0.0	-2.1
a	0.0	-1.1	0.4	-35.5
b	1.2	-1.9	-0.4	-7.7
c	-2.8	1.1	0.3	7.6
d	0.0	4.6	-3.6	21.6
K	0.0	0.0	-0.7	34.4
L	1.4	0.0	-0.6	-4.8
M	-4.6	0.4	4.8	0.0
ø	11.2	-1.1	-4.9	50.2

The procedure we are describing stops here.

It is quite possible to go on and sweep out again in the **147** direction, but we shall not do so. (Pressing on to eventual stability does NOT seem to be a good thing.) Sweeping once in each direction is enough to provide a considerable amount of resistance, as the reader may verify by sweeping out the analysis of table **46**. This is what we seek.

There are six orders in which we can sweep in the **147**, **KLM** and **abcd** directions. In general, each choice of order leads to a slightly different result. This is the price of less arithmetic.

Note that, when we go from analysis by means to the more resistant streamlined analysis, the q-typical square of the residuals has been raised from 2.8 to 3.2, while the typical midsquare has been reduced from 2.4 to 0.54, both changes being in the usual direction.

table 46 of chapter 14

Example of Analysis Suitable for Sweeping

	1	4	7	Ø
aK	12	-6	-6	6
aL	-6	3	3	-3
aM	-6	3	3	-3
bK	-4	2	2	-2
bL	2	-1	-1	1
bM	2	-1	-1	1
cK	-4	2	2	-2
cL	2	-1	-1	1
cM	2	-1	-1	1
dK	-4	2	2	-2
dL	2	-1	-1	1
dM	2	-1	-1	1
a	6	-3	-3	3
b	-2	1	1	-1
c	-2	1	1	-1
d	-2	1	1	-1
K	4	-2	-2	2
L	-2	1	1	-1
M	-2	1	1	-1
Ø	2	-1	-1	1